

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Somerville, Chris
Broun, Pierre
van de Loo, Frank

(ii) TITLE OF INVENTION: Production of Hydroxylated Fatty Acids in
Genetically Modified Plants

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: CUSHMAN DARBY & CUSHMAN, LLP
(B) STREET: 1100 NEW YORK AVENUE, N.W.
(C) CITY: WASHINGTON
(D) STATE: D.C.
(E) COUNTRY: USA
(F) ZIP: 20005-3918

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.44 KB storage
(B) COMPUTER: IBM PC
(C) OPERATING SYSTEM: DOS 5.0
(D) SOFTWARE: Word Perfect 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: unassigned
(B) FILING DATE: February 5, 1996
(C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Paul N. Kokulis
(B) REGISTRATION NUMBER: 16,773
(C) REFERENCE/DOCKET NUMBER: 1220/216442

(viii) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 202 861 3000
(B) TELEFAX: 202 822 0944

(2) INFORMATION FOR SEQ ID NO:1

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 543 nucleotides
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TATTGGCACC GGCGGCACCA TTCCAACAAT GGATCCCTAG AAAAGATGA AGTCTTGTC 60
CCACCTAAGA AAGCTGCAGT CANATGGTAT GTCAAATACC TCAACAACCC TCTTGGACGC 120
ATTCTGGTGT TAACAGTTCA GTTTATCCTC GGGTGGCCTT TGTATCTAGC CTTTAATGTA 180
TCAGGGTAGAC CTTATGATGG TTTCGCTTCA CATTCTTCC CTCATGCACC TATCTTAAG 240
GACCGTGAAC GTCTCCAGAT ATACATCTCA GATGCTGGTA TTCTAGCTGT CTGTTATGGT 300
CTTTACCGTT ACGCTGCTTC ACAAGGATTG ACTGCTATGA TCTGCGTCTA CGGAGTACCG 360
CTTTTGATAG TGAACCTTTT CCTTGCTTGT GTCACTTCT TGCAGCACAC TCATCCTTCA 420
TTACCTCACT ATGATTCAAC CGAGTGGAA TGGATTAGAG GAGCTTGTT TACGGTAGAC 480
AGAGACTATG GAATCTTGAA CAAGGTGTTT CACAACATAA CAGACACCCA CGTAGCACAC 540
CAC 543

(2) INFORMATION FOR SEQ ID NO:2

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 544 nucleotides
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TATAGGCACC GGAGGCACCA TTCCAACACA GGATCCCTCG AAAGAGATGA AGTATTTGTC 60
CCAAAGCAGA AATCCGCAAT CAAGTGGTAC GGCGAATACC TCAACAACCC TCCTGGTCGC 120
ATCATGATGT TAACTGTCCA GTTCGTCCTC GGATGGCCCT TGTACTTAGC CTTCAACGTT 180
TCTGGCAGAC CCTACAATGG TTTCGCTTCC CATTCTTCC CCAATGCTCC TATCTACAAC 240
GACCGTGAAC GCCTCCAGAT TTACATCTCT GATGCTGGTA TTCTAGCCGT CTGTTATGGT 300
CTTTACCGTT ACGCTGTTGC ACAAGGACTA GCCTCAATGA TCTGTCTAAA CGGAGTTCCG 360
CTTCTGATAG TTAACCTTTT CCTCGTCTTG ATCACTTAAC TACAACACAC TCACCCTGCG 420
TTGCCTCACT ATGATTCACTC AGAGTGGGAT TGGCTTAGAG GAGCTTAGC TACTGTAGAC 480
AGAGACTATG GAATCTTGAA CAAGGTGTTT CATAACATCA CAGACACCCA CGTCGCACAC 540

(2) INFORMATION FOR SEQ ID NO:3

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1740 nucleotides
 (B) TYPE: nucleotide
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAAGCTTT ATAAGAAGTT AGTTTCTCT GGTGACAGAG AAATTNTGTC AATTGGTAGT	60
GACAGTTGAA GCAACAGGAA CAACAAGGAT GGTTGGTGNT GATGCTGATG TGGTGATGTG	120
TTTATTTCATCA AATACTAAAT ACTACATTAC TTGTTGCTGC CTACTTCTCC TATTTCCCTCC	180
GCCACCCATT TTGGACCCAC GANCCTTCCA TTTAAACCCCT CTCTCGTGCT ATTCACCAGA	240
AGAGAAGCCA AGAGAGAGAG AGAGAGAATG TTCTGAGGAT CATTGTCTTC TTCATCGTTA	300
TTAACGTAAG TTTTTTTTGA CCACTCATAT CTAAAATCTA GTACATGCAA TAGATTAAATG	360
ACTGTTCCCTT CTTTGATAT TTTCAGCTTC TTGAATTCAA GATGGGTGCT GGTGGAAGAA	420
TAATGGTTAC CCCCTCTTCC AAGAAATCAG AAACTGAAGC CCTAAAACGT GGACCATGTG	480
AGAAACCACC ATTCACTGTT AAAGATCTGA AGAAAGCAAT CCCACAGCAT TGTTCAAGC	540
GCTCTATCCC TCGTTCTTTC TCCTACCTTC TCACAGATAT CACTTAGTT TCTTGCTTCT	600
ACTACGTTGC CACAAATTAC TTCTCTCTTC TTCCCTCAGCC TCTCTCTACT TACCTAGCTT	660
GGCCTCTCTA TTGGGTATGT CAAGGCTGTG TCTTAACCGG TATCTGGTCT ATTGGCCATG	720
AATGTGGTCA CCATGCATTC AGTGAATATC AATGGGTAGA TGACACTGTT GGTTTATCT	780
TCCATTCCCTT CCTTCTCGTC CCTTACTTCT CCTGGAAATA CAGTCATCGT CGTCACCATT	840
CCAACAATGG ATCTCTCGAG AAAGATGAAG TCTTGTCCC ACCGAAGAAA GCTGCAGTCA	900
AATGGTATGT TAAATACCTC AACAAACCTC TTGGACGCAT TCTGGTGTAA ACAGTTCACT	960
TTATCCTCGG GTGGCCTTTG TATCTAGCCT TTAATGTATC AGGTAGACCT TATGATGGTT	1020
TCGCTTCACA TTTCTCCCT CATGCACCTA TCTTTAAAGA CCGAGAACGC CTCCAGATAT	1080
ACATCTCAGA TGCTGGTATT CTAGCTGTCT GTTATGGTCT TTACCGTTAC GCTGCTTCAC	1140
AAGGATTGAC TGCTATGATC TGCGTCTATG GAGTACCGCT TTTGATAGTG AACTTTTCC	1200

TTGTCTTGGT AACTTTCTTG CAGCACACTC ATCCTTCGTT ACCTCATTAT GATTCAACCG	1260
AGTGGGAATG GATTAGAGGA GCTTTGGTTA CGGTAGACAG AGACTATGGA ATATTGAACA	1320
AGGTGTTCCA TAACATAACA GACACACATG TGGCTCATCA TCTCTTGCA ACTATACCGC	1380
ATTATAACGC AATGGAAGCT ACAGAGGCAGA TAAAGCCAAT ACTTGGTGAT TACTACCACT	1440
TCGATGGAAC ACCGTGGTAT GTGGCCATGT ATAGGGAAAGC AAAGGAGTGT CTCTATGTAG	1500
AACCGGATAC GGAACGTGGG AAGAAAGGTG TCTACTATTA CAACAATAAG <u>TTATGAGGCT</u>	1560
GATAGGGCGA GAGAAGTGCA ATTATCAATC TTCATTCCA TGTTTAGGT GTCTTGTAA	1620
AGAAGCTATG CTTTGTTCATAAATCTCAG AGTCCATNTA GTTGTGTTCT GGTGCATTT	1680
GCCTAGTTAT GTGGTGTGG AAGTTAGTGT TCAAAC TGCTGCTGTG CTGCCAGTG	1740
AAGAACAAAGT TTACGTGTTT AAAATACTCG GAACGAATTG ACCACAANAT ATCCAAAACC	1800
GGCTATCCGA ATTCCATATC CGAAAACCGG ATATCCAAAT TTCCAGAGTA CTTAG	1855

(2) INFORMATION FOR SEQ ID NO:4

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

10
Met Gly Ala Gly Gly Arg Ile Met Val Thr Pro Ser Ser Lys Lys Ser

20
Glu Thr Glu Ala Leu Lys Arg Gly Pro Cys Glu Lys Pro Pro Phe Thr

30
Val Lys Asp Leu Lys Lys Ala Ile Pro Gln His Cys Phe Lys Arg Ser

40
Ile Pro Arg Ser Phe Ser Tyr Leu Leu Thr Asp Ile Thr Leu Val Ser

50
Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro

60
Leu Ser Thr Tyr Leu Ala Trp Pro Leu Tyr Trp Val Cys Gln Gly Cys

70
90

100 Val Leu Thr Gly Ile Trp Val Ile Gly His Glu Cys Gly His His Ala
120 Phe Ser Asp Tyr Gln Trp Val Asp Asp Thr Val Gly Phe Ile Phe His
130 Ser Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg
150 His His Ser Asn Asn Gly Ser Leu Glu Lys Asp Glu Val Phe Val Pro
170 Pro Lys Lys Ala Ala Val Lys Trp Tyr Val Lys Tyr Leu Asn Asn Pro
180 Leu Gly Arg Ile Leu Val Leu Thr Val Gln Phe Ile Leu Gly Trp Pro
200 Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala
210 Ser His Phe Phe Pro His Ala Pro Ile Phe Lys Asp Arg Glu Arg Leu
230 Gln Ile Tyr Ile Ser Asp Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu
250 Tyr Arg Tyr Ala Ala Ser Gln Gly Leu Thr Ala Met Ile Cys Val Tyr
260 Gly Val Pro Leu Leu Ile Val Asn Phe Phe Leu Val Leu Val Thr Phe
280 Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Thr Glu Trp
300 Glu Trp Ile Arg Gly Ala Leu Val Thr Val Asp Arg Asp Tyr Gly Ile
310 Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His
330 Leu Phe Ala Thr Ile Pro His Tyr Asn Ala Met Glu Ala Thr Glu Ala
340 Ile Lys Pro Ile Leu Gly Asp Tyr Tyr His Phe Asp Gly Thr Pro Trp
360 Tyr Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Leu Tyr Val Glu Pro
370 Asp Thr Glu Arg Gly Lys Lys Gly Val Tyr Tyr Tyr Asn Asn Lys Leu

(2) INFORMATION FOR SEQ ID NO:5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 387 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Thr Ser Asn Asn Ser
20 Glu Lys Lys Gly Gly Ser Ser His Leu Lys Arg Ala Pro His Thr Lys
40 Pro Pro Phe Thr Leu Gly Asp Leu Lys Arg Ala Ile Pro Pro His Cys
50 Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Tyr Val Ala Tyr Asp Val
70 Cys Leu Ser Phe Leu Phe Tyr Ser Ile Ala Thr Asn Phe Phe Pro Tyr
90 Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu Phe
100 Gln Gly Cys Ile Leu Thr Gly Leu Trp Val Ile Gly His Glu Cys Gly
120 His His Ala Phe Ser Glu Tyr Gln Leu Ala Asp Asp Ile Val Gly Leu
130 Ile Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser
150 His Arg Arg His His Ser Asn Ile Gly Ser Leu Glu Arg Asp Glu Val
170 Phe Val Pro Lys Ser Lys Ser Ile Ser Trp Tyr Ser Lys Tyr Ser
180 Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Ala Ala Thr Leu Leu Leu
200 Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp
210 Arg Phe Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile Phe Ser Glu Arg
220

230 240
Glu Arg Leu Gln Ile Tyr Ile Ala Asp Leu Gly Ile Phe Ala Thr Thr
250
Phe Val Leu Tyr Gln Ala Thr Met Ala Lys Gly Leu Ala Trp Val Met
260 270
Arg Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Cys Phe Leu Val Met
280
Ile Thr Tyr Leu Gln His Thr His Pro Ala Ile Pro Arg Tyr Gly Ser
290 300
Ser Glu Trp Asp Trp Leu Arg Gly Ala Met Val Thr Val Asp Arg Asp
310 320
Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His Val
330
Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu Ala
340 350
Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp Gly
360
Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu Phe
370 380
Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr Arg
Asn Lys Tyr

(2) INFORMATION FOR SEQ ID NO:6

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 383 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

10
Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
20 30
Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser

40
Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
50 60
Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser
70 80
Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
90
Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
100 110
Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
120
Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
130 140
Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
150 160
His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
170
Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
180 190
Gly Arg Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Trp Pro Leu
200
Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys
210 220
His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
230 240
Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr
250
Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly
260 270
Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu
280
Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp
290 300
Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
310 320
Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu

330
Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile
340 350
Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr
360
Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp
370 380
Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu

(2) INFORMATION FOR SEQ ID NO:7

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 384 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

10
Met Gly Ala Gly Gly Arg Met Gln Val Ser Pro Pro Ser Lys Lys Ser
20 30
Glu Thr Asp Asn Ile Lys Arg Val Pro Cys Glu Thr Pro Pro Phe Thr
40
Val Gly Glu Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
50 60
Ile Pro Arg Ser Phe Ser His Leu Ile Trp Asp Ile Ile Ile Ala Ser
70 80
Cys Phe Tyr Tyr Val Ala Thr Thr Tyr Phe Pro Leu Leu Pro Asn Pro
90
Leu Ser Tyr Phe Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
100 110
Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His Ala Ala Phe
120
Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
130 140
Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His

09885489-052102
His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Arg
Arg Ser Gln Thr Ser Ser Gly Thr Ala Ser Thr Ser Thr Thr Phe Gly
Arg Thr Val Met Leu Thr Val Gln Phe Thr Leu Gly Trp Pro Leu Tyr
Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Gly Phe Ala Cys
His Phe His Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
Ile Tyr Ile Ser Asp Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu Leu
Pro Tyr Ala Ala Val Gln Gly Val Ala Ser Met Val Cys Phe Leu Arg
Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu Ile Thr Tyr Leu
Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp
Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
Asn Gln Gly Phe His Asn Ile Thr Asp Thr His Glu Ala His His Leu
Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile
Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Gly Thr Pro Val Val
Lys Ala Met Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp
Arg Gln Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu Xaa

(2) INFORMATION FOR SEQ ID NO:8

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 309 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

10
Ser Leu Leu Thr Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Phe Ala
20 30
Phe Ile Phe Tyr Ile Ala Thr Thr Tyr Phe His Leu Leu Pro Gln Pro
40
Phe Ser Leu Ile Ala Trp Pro Ile Tyr Trp Val Leu Gln Gly Cys Leu
50 60
Leu Thr Arg Val Cys Gly His His Ala Phe Ser Lys Tyr Gln Trp Val
70 80
Asp Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val Pro Tyr
90
Phe Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr Gly Ser
100 110
Leu Asp Arg Asp Glu Arg Val Lys Val Ala Trp Phe Ser Lys Tyr Leu
120
Asn Asn Pro Leu Gly Arg Ala Val Ser Leu Leu Val Thr Leu Thr Ile
130 140
Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp
150 160
Ser Phe Ala Ser His Tyr His Pro Tyr Arg Val Arg Leu Leu Ile Tyr
170
Val Ser Asp Val Ala Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val
180 190
Ala Thr Leu Lys Gly Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro
200
Leu Leu Ile Val Asn Gly Phe Leu Val Thr Ile Thr Tyr Leu Arg Val
210 220
His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala Thr
230 240
Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His His Ile Thr
250
Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His
260 270
Leu Arg Val Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Asp Thr

280
Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Arg Glu Cys Leu Tyr Val
290 300
Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly Val Tyr Trp Tyr Arg Asn
Lys Tyr Leu Arg Val

(2) INFORMATION FOR SEQ ID NO:9

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 302 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

10
Phe Ser Tyr Val Val Tyr Asp Leu Thr Ile Ala Phe Cys Leu Tyr Tyr
20 30
Val Ala Thr His Tyr Phe His Leu Leu Pro Gly Pro Leu Ser Phe Arg
40
Gly Met Ala Ile Tyr Trp Ala Val Gln Gly Cys Ile Leu Thr Gly Val
50 60
Trp Val Val Ala Phe Ser Asp Tyr Gln Leu Leu Asp Asp Ile Val Gly
70 80
Leu Ile Leu His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr
90
Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu
100 110
Val Phe Val Pro Lys Val Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg
120
Val Leu Thr Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu
130 140
Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr
150 160
Asp Pro Tyr Gly Pro Ile Tyr Ser Val Ile Ser Asp Ala Gly Val Leu
170
Ala Val Val Tyr Gly Leu Phe Arg Leu Ala Met Ala Lys Gly Leu Ala

180 190
Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu Val Val Asn Gly Phe
200
Leu Val Leu Ile Thr Phe Leu Gln His Thr His Val Ser Glu Trp Asp
210 220
Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
230 240
Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
250
Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Val Glu Tyr
260 270
Tyr Arg Phe Asp Glu Thr Pro Phe Val Lys Ala Met Trp Arg Glu Ala
280
Arg Glu Cys Ile Tyr Val Glu Pro Asp Gln Ser Thr Glu Ser Lys Gly
290 300
Val Phe Trp Tyr Asn Asn Lys Leu Ala Met Glu Ala Thr Val

(2) INFORMATION FOR SEQ ID NO:10

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

10
Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
20 30
Gln Leu Ala Arg Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val
40
Glu Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
50 60
His Cys Phe Glu Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His
70 80
Asp Leu Val Ile Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile
90
Pro Ala Leu Pro Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp

100 Ile Ala Gln Gly Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Val Val
120 Gly Leu Val Leu His Ser Ser Leu Met Val Pro Tyr Phe Ser Trp Lys
130 Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp
150 Glu Val Phe Val Pro Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro
170 Tyr Val Tyr Asn Asn Pro Val Gly Arg Val Val His Ile Val Val Gln
180 Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg
200 Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr
210 Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Val
230 Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Phe Gly Val Trp
250 Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp
260 Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu Pro His
280 Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met
290 Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
310 Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala
330 Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His
340 Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu
360 Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn
370 Lys Lys Phe Xaa

(2) INFORMATION FOR SEQ ID NO:11

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 224 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

10
Trp Val Met Ala His Asp Cys Gly His His Ala Phe Ser Asp Tyr Gln
20
Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val
40
Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr
50
Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Lys Ser Ser
70
Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met
90
Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe
100
Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro
120
Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser
130
Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile
150
Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu
170
Val Val Asn Ser Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His
180
Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly
200
Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe
210
His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Phe Thr Met Pro

(2) INFORMATION FOR SEQ ID NO:12

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCTCTTTGT GCGCTCATTC

20

(2) INFORMATION FOR SEQ ID NO:13

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGTACCAAGA AAACGCCTTG

20

(2) INFORMATION FOR SEQ ID NO:14

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TAYWSNCAYM GNMGNCA

20

(2) INFORMATION FOR SEQ ID NO:15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 nucleotides
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

RTGRTGNGCN ACRTGNGTRT C

21

09885189 - 062104